# 一种面向 FAST PB 量级脉冲星数据处理加速方法及系统*

张辉 [1,3,4]，谢晓尧 [1,4]，李菂 [2,5]，刘志杰 [1,4]，王培 [2]，于徐红 [1,4]，游善平 [1,4]，许余云 [6]，姜家涛 [1,3,4]

（1.贵州师范大学省信息与计算科学重点实验室，贵阳 550001；2.中国科学院国家天文台，北京 100012；3.贵州师范大学数学科学学院，贵阳 550001；4.FAST 早期科学数据中心，贵阳 550001；5.中国科学院大学，北京 100049; 6.贵州水利水电职业技术学院管理工程分院，贵阳 551416)

**摘 要：**500 米口径球面射电望远镜 FAST 已投入科学运行，漂移扫描巡天采集数据量已超过 1PB，预计每年新增至少 5PB。现有数据处理软件如 PRESTO、SIGPROC 等已无法满足 PB 量级数据的快速处理要求。提出了一种基于 PRESTO 的分布式并行计算方法，整合利用数据库技术和异地异构计算资源，构建了一套命名 Craber 的计算加速系统。整体系统由 FAST 早期科学数据中心与国家天文台团队共同设计和实现。启用 Craber 子网计算集群 D 中 55 个计算节点，应用 Parkes 多波束巡天数据集和 FAST 漂移扫描数据验证了系统流程和搜索数据库。单个 100MB Parkes 巡天数据文件平均耗时 36 秒，单个 128MB FAST 巡天数据文件平均耗时 22 秒。该系统目前已实际参与 FAST 数据处理并产生数十颗脉冲星发现，有效帮助 FAST 加速数据处理和扩大新样本数量。

**关键词：**脉冲星；加速系统；搜索数据库；数据处理；500 米口径球面射电望远镜
**图分类号：**P162 **文献标识码：**A **文章编号：**

1967年，剑桥大学在读博士生乔瑟琳·贝尔(Jocelyn Bell)和其导师等人在狐狸星座发现了人类历史上首颗射电脉冲星PSR B1919+21[1]。这项20世纪重大的天文学发现于1974年12月获得了诺贝尔物理学奖，以表彰赖尔（Martin Ryle）和休伊什（Antony Hewish）教授在射电天体物理学领域的开创性研究。至今已知探测发现脉冲星约3000颗，它们的自行速度每秒可达数百公里[2]。脉冲星极端物理性质是地面人工实验室无法实现的，性质包括：强磁场、高密度、高能辐射、超强引力场、伴随恒星演化和超新星爆发过程等。脉冲星的发现为核物理、粒子物理、天体物理、爱因斯坦相对论理论和宇宙学的检验提供天然理想实验室。对其深入研究，有希望获得许多重大物理学问题的答案。如：它自转周期极其稳定，具有精准的时钟信号，可为引力波探测、航天器导航等重大科学技术应用提供理想工具。

脉冲星独特的物理性质和科学价值使其很快成为天体物理学最活跃的研究领域之一。为更深入研究这类天体，大量望远镜时间被投入开展巡天项目，以期发现更多样本数据。过去，受望远镜设备、数据采集装置、计算搜索理论和技术等因素限制，单个巡天项目采集到的观测数据总量通常介于GB至TB之间。例如，1997-2003年期间澳大利亚Parkes望远镜采用13波束接收机巡天（Parkes Multi-beam Pulsar Survey, PMPS）[3]，采集到49700余个PSRFITS格式数据文件[4]，数据总量约5TB，探测发现新脉冲星超过1122颗。2012年Nançay望远镜开展的SPAN512项目，采集数据总量约50TB。而2010-2012年开展的“高时间分辨宇宙（HTRU）”系列巡天项目，因同时提高了时间和频率分辨率，每个Beam文件达16GB，采集数据总量约1PB。如今，500米口径球面射电望远镜FAST已投入使用，可探测频率覆盖范围70MHz至

---

*基金项目：国家自然科学基金(U1831131，U1631132，U1731238，11743002)；中国科学院天文大科学研究中心 FAST 重大成果培育项目(FAST[2019sr04])；2017 贵州省科技计划(201742920352110100)；贵州师范大学研究生创新基金(研创 201528)资助.
收稿日期：2020-05-06; 修订日期：2020-05-19
作者简介：张辉(1985-)，男，博士研究生. 研究方向：大数据分析与处理. Email：zhanghui@gznu.edu.cn
通讯作者：谢晓尧，男，教授. 博士研究生导师. 研究方向：智能计算、大数据分析与处理. Email：xyx@gznu.edu.cn

3GHz，它灵敏度更高，综合探测能力更强，具备19波束同时扫描巡天能力，配备更高时间、频率分辨率的超宽带接收机，单波束数据采样频率由原来的1024Hz拓展至4096Hz。若19个波束同时工作，以100 $\mu s$ 为时间采样分辨率、频率带宽400MHz，采用8 bit记录数据，日采集数据量可达100TB。2017年7月-2018年5月期间，FAST使用超宽带接收机进行夜间漂移扫描( Commensal Radio Astronomy FAST survey, CRAFTS)，记录0-1GHz，1-2GHz两个频段数据，接收到观测数据超过1400TB，其中，超宽带数据585TB，19波束巡天数据851TB，加上数据分析产生的中间数据，已超过1.5PB容量，数据存于FAST早期科学数据中心集群磁盘阵列或磁带库中。预计FAST在射电脉冲星科学领域总采集脉冲星数据量在10PB-100PB规模，FAST巡天数据已迈入PB时代[5]。

FAST面临大规模数据地传输、存储、快速处理和结果数据管理等实际问题。针对FAST PB量级数据分析处理加速问题，本文提出了一种基于PRESTO[6]的分布式并行计算方法，并据此研发了一套并行计算加速系统。系统命名为Craber，以纪念中国科学家记载Crab脉冲星发现和脉冲星科学工作者开拓创新的科研精神。

## 1 搜索流程、软件及加速分析

### 1.1 搜索流程和软件

从观测数据中搜索脉冲星涉及多个必要步骤，包括：去干扰、消色散、时频信号转换、周期信号搜索、数据折叠和候选体识别等。以计算机处理数据的视角看，搜索流程可概括为：(1)获取数据文件。计算节点读取一个观测原文件。(2)分析处理数据。按照预先设定的参数空间和搜索策略，执行分析指令。(3)生成结果，生成可供识别的候选体集合。(4)读取下一个待处理文件，并重复执行(1)、(2)、(3)步，详细的搜索流程和技术理论参见 Lorimer 专著[7]。潘之辰等在脉冲星搜索加速方面也做过尝试，取得了一定加速效果[8]。目前，针对脉冲星数据处理较好的软件有：PRESTO、SIGPROC 和基于 GPU 技术的 peasoup 等。但存在如下几方面不足。

1. 软件并行程度不够。在一般的搜索中，消色散方法主要有两大类：相干消色散和非相干消色散[9]。需尝试数百个甚至上千个色散值，计算量便会成百上千倍增加。现有软件因开发历史较早，基于当时较小的数据规模，搜索流程中大多采用分步串行、单机计算方式。

2. 计算结果整理、统计和归档能力弱。计算参数、中间数据、结果数据、日志数据等缺乏科学化管理，数据归档同步不及时、数据分享和统计分析困难。

3. 计算资源整合能力差。观测数据以FITS数据格式存储，搜索中伴随产生临时众多文件，且多为小文件，带来频繁读写磁盘，计算任务具有数据密集型和计算密集型双重特点。现有搜索软件缺乏整合计算资源的能力，特别是不具备整合异地异构计算资源的方法。

### 1.2 并行计算分析

针对现有搜索软件存在的前述问题，结合 FAST 数据搜索特点和实际应用，对计算可行性分析如下。

1. 任务可并行性。从搜索流程分析，不同观测文件间不存在数据耦合关系，且处理过程中也保持数据相互独立性。因此，全体待处理观测文件构成一个任务集合，单个观测文件为其中一个子计算任务。

2. 计算节点间可并行性。通过网络技术，将单个计算节点整合起来，特别是将异地计算节点整合起来，形成一定算力的集群系统，各单节点上并行处理任务。

3. 软件程序可并行性。根据 PRESTO 模块化程序设计和分步串行计算特点，经分析发现它具备改并行运算的可能。可通过在各节点上部署优化过的 PRESTO，并行调度其子程

序进行计算。

4. 搜索数据库协同计算和数据管理。异地计算资源协同计算和多任务并行处理面临数据一致性和归档问题，可通过建立搜索数据库系统协同管理和结果回收。

### 1.3 FAST 早期科学数据中心计算资源

FAST 早期科学数据中心由贵州师范大学与国家天文台联合建设，目标是解决 FAST 数据存储、分析处理和开展前沿科学研究。早期科学数据中心至 FAST 台址两地数据中心通过 2GB 专网连接，横跨州、县、镇 20 余中转节点，传输距离 300 多公里，采用波分复用和双路通信容错技术保障数据传输的可用性。数据中心致力于改善软件可用性的同时，也设计、定制和组装满足 FAST 未来一段时期计算需求的计算资源。

如表 1，数据中心已分步建成 A、B、C、D 四类不同子网计算集群，配备 1.5PB 磁盘存储系统和 1PB 容量磁带库，部署在数据中心宝山、花溪两地机房。计算集群采用 CPU+GPU 异构并行计算体系架构，包含 1384 个 CPU 核心，288 块 GPU 显卡，共计 1384 个 CPU 核心和 955392 多个 GPU 核心。并对台式 PC 机进行改造，扩展显卡插槽，加装 GPU 显卡、大容量磁盘和内存资源等，强化单节点计算能力,共同组成并行化计算集群，由任务管理节点统一调度。极大地解决 FAST 当前一段时间的数据分析任务。

表 1 FAST 早期科学数据中心 CPU+GPU 异构系统计算集群算力统计

Table 1 Statistics for computational power of computing cluster in FAST Early Science Data Center

| 子网集群 | 计算节点数目(个) | CPU 数目（块） | CPU 核心数(个) | GPU 数目（块） | GPU 核心数(个) |
|---|---|---|---|---|---|
| 子网集群 A | 10 | 20 | 440 | 80 | 286720 |
| 子网集群 B | 60 | 60 | 480 | 120 | 430080 |
| 子网集群 C | 14 | 14 | 224 | 28 | 100352 |
| 子网集群 D | 60 | 60 | 240 | 60 | 138240 |
| 合计 | 144 | 154 | 1384 | 288 | 955392 |

## 2 系统设计

### 2.1 设计思路和方法

根据数据中心已有的软硬件资源，整合数据库、搜索软件和网络，解决如何调度、如何管理、如何抗错等问题，系统设计思路和方法如下。

1. 优化 PRESTO 程序套件，开发一套并行计算任务调度系统，调度部署在各节点上的 PRESTO 程序各分步搜索指令，并行的分析处理子任务。

2. 整合利用 FAST 早期数据中心位于宝山、花溪两校区子网计算集群，构建一个对外透明的分布式并行计算集群。

3. 以单个观测数据文件划分计算任务单元，建立任务队列，经任务调度分发程序和数据库协同具体搜索工作。

4. 建立面向搜索过程和数据管理的专门数据库，进行任务分发、计算协同、日志数据和结果数据回收管理，解决异地计算资源参与计算时的数据一致性和归档问题，供后续 AI 程序开展候选体识别和筛选。

### 2.2 功能节点划分

如图1，Craber并行加速计算系统按照功能不同，逻辑上划分为以下几个组成部分：一个数据源节点，用于存放FITS观测数据文件；一个任务控制节点，用于计算任务的协调控制；N个并行计算节点，负责具体计算搜索业务；一个数据库节点，承担数据管理和协同计算；一个非结构化结果数据节点，负责存放过程数据；一个WEB应用节点，提供配置交互。各节点间通过高速网络连接进行通信和数据交换，调度系统采用C/S架构实现。

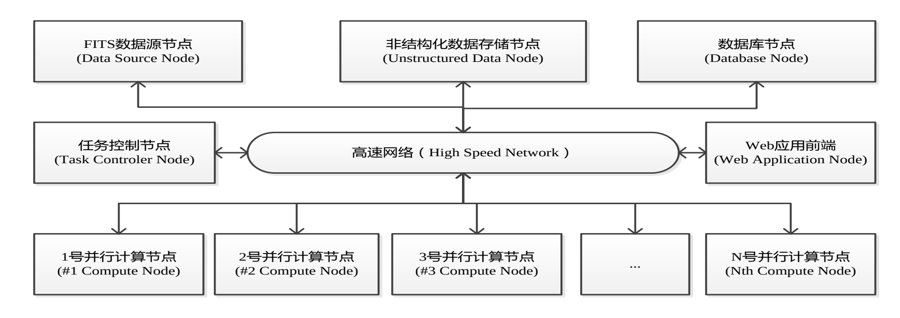



图 1 Craber 加速计算系统功能结构图

Fig.1 Structure diagram of the Craber parallel accelerated computing system

## 2.3 计算节点与其他节点间指令和数据传递过程

计算节点主动与数据库和任务控制节点通信和获取计算任务。任务节点负责从数据库任务队列中获取计算任务参数，分发给计算节点，如数据文件地址、计算参数、处理指令集。计算节点作为客户端，通过其上的配置文件获取参数信息，与服务器通信后，进入计算状态，执行 PRESTO 套件程序指令，进行脉冲星数据的计算工作，并将计算结果提交回任务控制服务器和计算结果文件服务器。如表 2，节点上配置文件规定了此节点用于去干扰、消色散、周期搜索、数据折叠等步骤的线程分配方案，各步等待时间以及数据同步情况。

表 2 配置文件参数列表说明

Table2 Configuration file parameter list and description

| Key word | Description | Key word | Description |
|---|---|---|---|
| updatesleeptime | 写数据库操作线程等待时间 | tcpserverip | 服务器 IP 地址 |
| sendtodatabase | 判读是否需要写入数据库 | tcpserverport | 通信连接端口 |
| threadnum | 各计算步骤线程队列可用线程数 | prepsubbandmaxstep | 消色散 DM 步长值 |
| sleeptime | 各计算步骤工作线程等待时间 | cachefile | 客户端计算 cache 目录 |
| readoncetimes | 广播间隔时间 | sourcefile | FITS 源文件目录地址 |
| readoncenodenum | 一次广播的客户端计算节点数 | savefile | 结果保存目录地址 |

## 2.4 计算节点内部数据处理和交互过程

单个节点的计算能力体现为可用的线程数量。任务分配服务器分配数个任务给计算节点，广播计算参数。如图 2，计算节点根据任务参数，从数据节点中读取对应观测数据文件并作处理。计算节点依据线程分配方案和数据库中任务 ID 值，并行的执行图中 1-6 步流程，数据同步由数据库和任务控制节点完成。

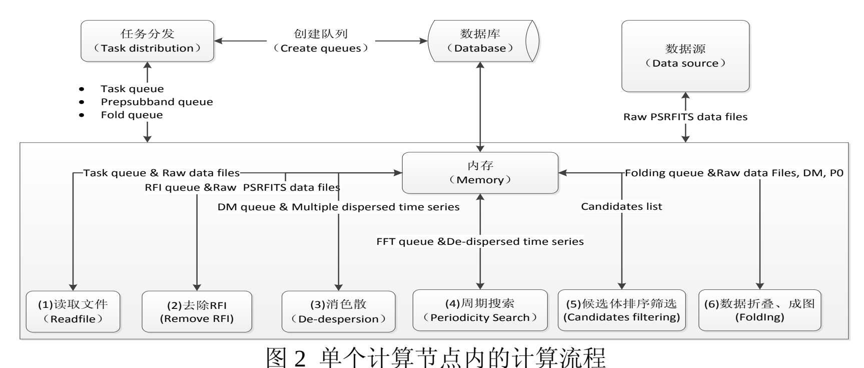



图 2 单个计算节点内的计算流程

Fig.2 Search process of one single compute node in the cluster

如图 3，单个计算节点 GPU/CPU 异构系统开展单个任务计算数据处理流程。在每个计算节点上通过指定 GPU 核心分配方案，分别用于 PRESTO 子程序并行的进行消除干扰(RFIFIND)、消除色散(PRESUBBAND)、周期搜索、数据折叠等步骤。因消色散耗时相对较长，系统指定较多核心数专门进行消色散，较少核心数进行周期搜索、数据折叠等步骤，进行流水化作业。

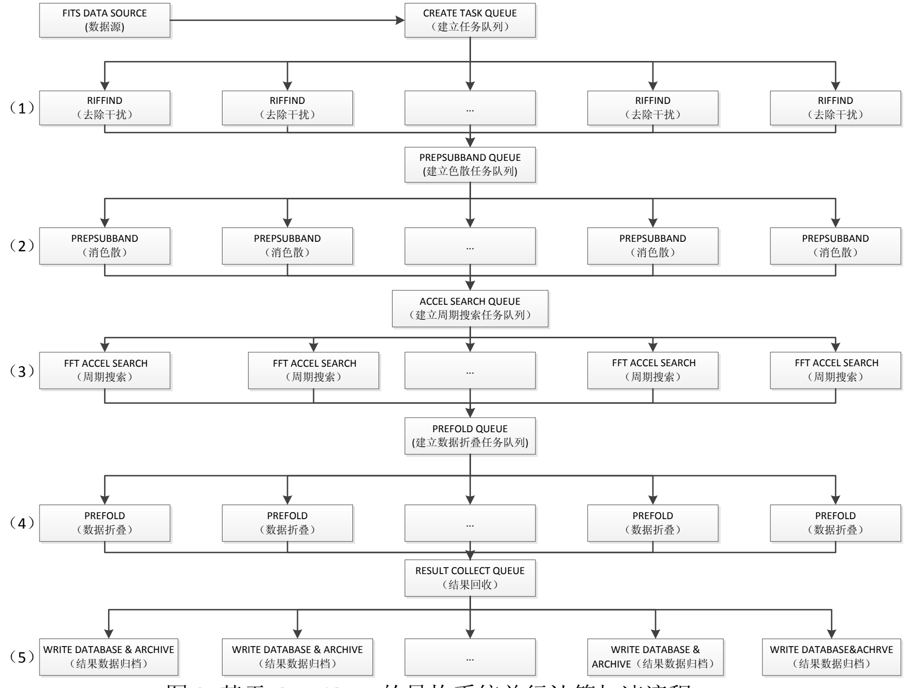



图 3 基于 GPU/CPU 的异构系统并行计算加速流程

Fig.3 GPU+CPU-based parallel computing acceleration process for heterogeneous systems

### 2.5 搜索数据库

Craber 系统的任务状态和任务数据的维护通过数据库来实现。搜索数据库解决的问题包括：1）在分布式计算任务中，保证调度的任务数据的一致性，提高容错能力；2）为原始数据、处理日志、结果整理和统计提供基础平台；3）为后续 AI 候选体识别筛选提供数据支持和辅助工具。设计了 FITS 观测数据文件信息表、任务执行时间状态信息表、任务指令参数表等 10 余个不同用途的数据库表。其中，文件信息表记录 FITS 文件的编号、文件名、文件路径、格式、观测设备和处理状态等。通过调用 PRESTO 中 readfile 指令获取数据集中全体待计算 FITS 头文件信息，并写入数据库任务指令参数表记录了一个任务 ID 对应的计算参数，以及 PRESTO 子程序指令：rfifind，prepsubband，realfft，accelsearch，prepfold 步骤编号，用来表示 prepsubband 的子步骤。搜索数据库详细内容将在其他文章专门介绍。

## 3 系统实施与结果分析

### 3.1 计算实验环境

如图 4，Craber 集成一个大容量数据文件服务器，存放经预处理后的 FITS 数据文件，为方便计算节点快速读取观测文件，数据源利用网络文件系统（Network File System）来实现管理，节点内使用临时文件系统(Temporary File System)提高 I/O 能力，部署一个任务控制服务器，用于计算任务的协调控制、任务分发和参数广播，数据中心 4 个计算子网分布式部署，负责具体计算搜索业务。搜索数据库采用大型关系型数据库 ORACLE 11g，进行结果数据管理和计算参数设定。一个文件服务器存放中间数据文件及候选体判别图。一个前端 WEB 应用程序，进行作业期间的实时监控。

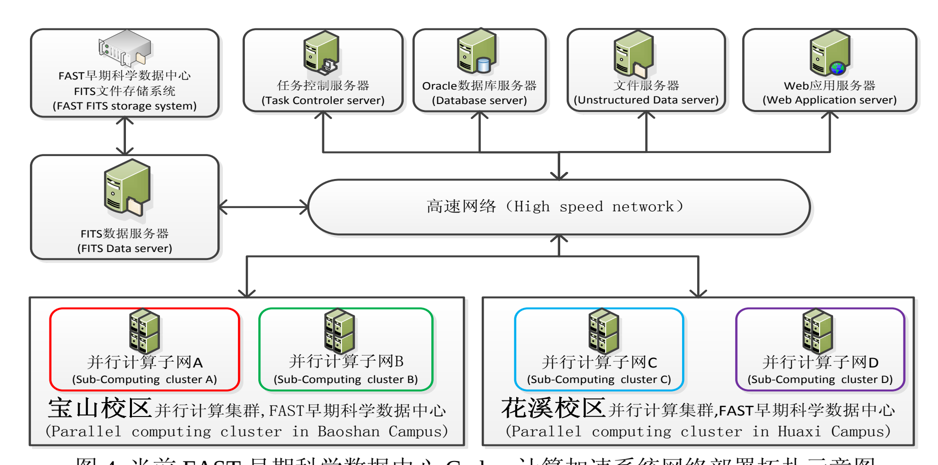



图 4 当前 FAST 早期科学数据中心 Craber 计算加速系统网络部署拓扑示意图

Fig.4 Network deployment topology of Craber acceleration system of FAST Early Science Data Center

## 3.2 测试数据

如表3，Parkes 多波束巡天项目PMPS采用96个频率通道，子带宽3MHz，积分时间2100秒，单个文件大小100MB。FAST望远镜 CRAFTS数据集经过频域分割数据预处理，频率范围269.125-396.875MHz，频道数512，带宽128MHz，单波束扫描时间26.2144s，单个文件大小128MB。

表3 测试数据文件样例

Table 3 Two sample test files from PMPS and FAST

| 关键字（Key word） | Parkes 望远镜数据文件（PMPS data file） | Parkes 望远镜数据文件（FAST data file） |
|---|---|---|
| File Name | PM0001_00731.sf | FL20170924_0QT30.fits |
| Pulsar | J1345-6115 | J2000-1234 |
| Number of Channels | 96 | 512 |
| Width of Channels | 3 MHz | 0.25 MHz |
| Band | 1231.5-1516.5 MHz | 269.125-396.875MHz |
| Time per file | 2100.224s | 26.2144s |
| Size of file | ~100MB | ~128MB |
| Time cost | 36s | 22s |

## 3.3 实验结果

如图5，随机选取了Parkes 400个PMPS数据文件，消色散DM步数设为768。以单线程单机处理，单个文件耗时约6000秒(100分钟)。采用8线程OMP处理，单个文件平均速度3530秒(约60分钟)。启用Craber子网计算集群D中55个计算节点，每个节点配置Intel酷睿i7处理器、32GB内存和1TB硬盘。总耗时约14550秒，平均36s处理一个数据文件，即集群整体吞吐速度是36s。

FAST 超宽带漂移扫描巡天数据分析测试，单个128MB数据文件平均处理时间22秒。2017年10月年至今参与处理FAST数据文件66000多个，相关结果数据统一存处于系统集成的搜索数据库和文件服务器中。FAST通过多套独立的脉冲星搜索流程，已确认发现114颗脉冲星，该系统已辅助发现数十颗新脉冲星[10][11]。辅助发现第1颗脉冲星J1859-0131（FP1），

自转周期1.83秒，距离地球约1.6万光年，第2颗脉冲星编号J1931-01（FP2），自转周期0.59秒，距离地球约4100光年，两颗脉冲星由FAST在南天银道面通过漂移扫描采集数据发现[12][13]。

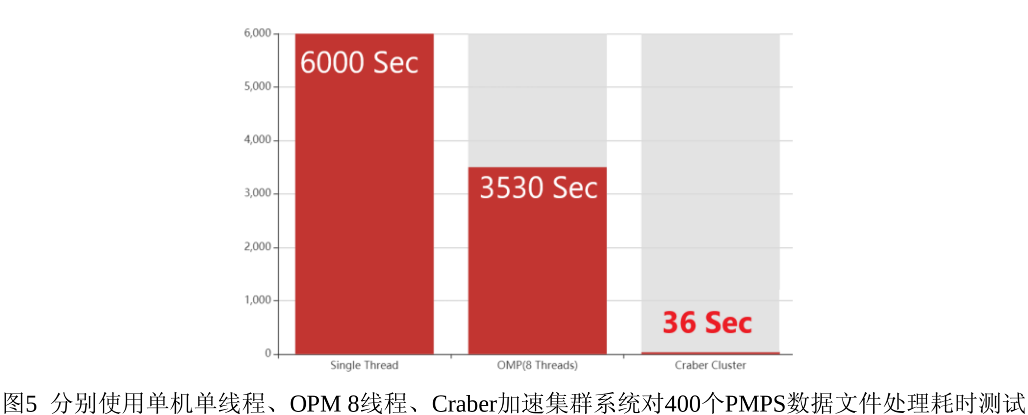



图5 分别使用单机单线程、OPM 8线程、Craber加速集群系统对400个PMPS数据文件处理耗时测试

Fig.5 Performance test for the processing time of 400 PMPs data files by using stand-alone single thread, OPM 8 thread and craber acceleration cluster system respectively

## 4 总结与展望

Craber计算加速系统包含154个CPU，288块GPU显卡,为FAST脉冲星搜索项目的数据处理提供了有力保障，集成的面向搜索过程的数据库系统在协同计算任务中，保证调度的任务数据的一致性，提高容错能力，并为原始数据、处理日志、结果整理和统计提供基础平台，同时，也为下一步AI程序开展学习训练、候选体识别提供数据支持，大幅减少结果数据整理统计和管理工作量。系统应用于Parkes巡天数据处理，单个100MB数据文件平均耗时36s，应用于FAST脉冲星漂移扫描巡天数据处理，单个数据文件平均耗时22秒。FAST使用多套独立的脉冲星搜索流程，确认发现114颗脉冲星，其中，该系统已辅助发现数十颗新脉冲星。系统可进适当参数调整被用于单脉冲(RRATs、FRBs)搜索，未来将应用FAST搜寻模式观测项目数据处理。随着5G技术成熟和硬件成本进一步降低，大规模数据的异地协同计算将变得更易操作。下一步计划持续向Craber集群系统增添更多计算节点，优化搜索数据库功能结构和计算节点中CPU/GPU调度方案，以应对FAST今后PB数据集的处理，改善数据中心存储压力，辅助探索发现更多新脉冲星样本。

**致谢**：感谢 Scott Ransom 提供的 PRESTO 程序套件，以及来自 Parkes 望远镜的数据支持。同时，特别感谢贵阳学院张正东博士、澳大利亚 CSRIO 王晨教授、代实博士以及网易的张翔工程师对本文的建设性建议和帮助。

---

**参考文献**

[1] Hewish A, Bell S J, Pilkington J D H, et al. Observation of a Rapidly Pulsating Radio Source[J]. Nature 217, 709–713 (1968).

[2] J.Khoo, G.Hobbs, R.N.Manchester, D.Miller, et al. Parkes 脉冲星数据库使用方法(英文)[J].天文研究与技术,2012,9(03):229-236.

[3] Manchester R.N., Lyne A.G., Camilo F., et al. The Parkes multi-beam pulsar survey – I. Observing and data analysis systems, discovery and timing of 100 pulsars. 2001, 328(1):17-35.

[4] https://casper.ssl.berkeley.edu/wiki/ROACH-2 Revision 2.

[5] Di Li, Zhichen Pan. The Five‑hundred‑meter Aperture Spherical Radio Telescope project[J].Radio Science. 2016, 51(7):1060-1064.

[6] http://www.cv.nrao.edu/~sransom/presto/.
[7] Lorimer D R , Kramer M . Handbook of Pulsar Astronomy[D]. 2012.
[8] 潘之辰，钱磊，岳友岭. 脉冲星搜索技术及 FAST 望远镜脉冲星搜索展望[J]. 天文研究与技术，2017， 14(01):8-16.
[9] 黄玉祥，汪敏，郝龙飞，等. 脉冲星信号相干消色散与非相干消色散的比较研究[J]. 天文研究与技术,2019,16(01):16-24.
[10] https://crafts.bao.ac.cn/pulsar/.
[11] A. D. Cameron, D. Li, G. Hobbs, et al. An in-depth investigation of 11 pulsars discovered by FAST[J]. MNRAS, arXiv: 2005.09171. DOI:10.1093/mnras/staa1328.
[12] Lei Qian, ZhiChen Pan, Di Li, et al. The first pulsar discovered by FAST[J]. Science China (Physics, Mechanics & Astronomy), 2019,62(05):75-78.
[13] http://www.bao.ac.cn/xwzx/zyxw/201710/t20171010_4871382.html.

# A Data Processing Acceleration Method and System for FAST Petabyte Pulsar Data Processing

Zhang Hui[1,3,4], Xie Xiaoyao[1], Li Di[2,5], Liu Zhijie[1,4], Wang Pei[2], Yu Xuhong[1,4], You Shanping[1,4], Xu Yuyun[6], Jiang Jiatao[1,3,4]

(1. Guizhou Normal University, Key Laboratory of Information and Computing Science, Guiyang, 550001, China; Email: xyx@gznu.edu.cn; 2. National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012, China; 3. Guizhou Normal University, School of Mathematics and Sciences, Guiyang, 550001, China; 4.FAST Early Science Data Center, Guiyang, 550001, China; 5. University of Chinese Academy of Sciences, Beijing 100049, China; 6.Guizhou Vocational and Technical College of Water Resources and Hydropower, Institute of Management Engineering, Guiyang, 551416, China)

**Abstract:** Five-hundred-meter Aperture Spherical Telescope has been put into operation. Data collected by drift scan pulsar survey has exceeded 1PB, and it is expected to increase at least by 5 PB per year. Existing data processing software, such as PRESTO, SIGPROC, etc., couldn't meet the real-time data analysis and management requirements. How to efficiently process PB volume of data has become a new challenge in the field of radio astronomy. In order to deal with the problem of PB data analysis and result data management encountered by FAST, we designed a PRESTO-based, distributed parallel computing method, which integrates network technology, database and cross-regional hardware computing resources. We also developed a acceleration computing cluster named Craber, which was designed and implemented by a team from Guizhou Normal University and the National Astronomical Observatories. The work flow and search-oriented database of the acceleration system are tested with data sets from the Parkes Multibeam Pulsar Survey and FAST Drift Scan Pulsar Survey. A 100MB Parkes data file takes ~36 seconds by 55 computing nodes. While applied to the FAST drift scan pulsar survey, a 128MB data file costs ~22 seconds. Craber has participated in the data processing of more than 66,000 FAST data files, helped FAST detect more than 141 high-quality candidates, 114 of them have been confirmed to be new pulsars. All result data stored in the integrated Oracle database or file server for further candidates selection with AI. It helped FAST speed up the data processing and find number of new pulsars.

**Key words:** Pulsar; Acceleration system; Database; Data processing; FAST